

# A Boolean Network Inference from Time-Series Gene Expression Data Using a Genetic Algorithm

Shohag Barman and Yung-Keun Kwon\*

School of IT Convergence, University of Ulsan, 93 Daehak-ro,  
Nam-gu, Ulsan, Republic of Korea

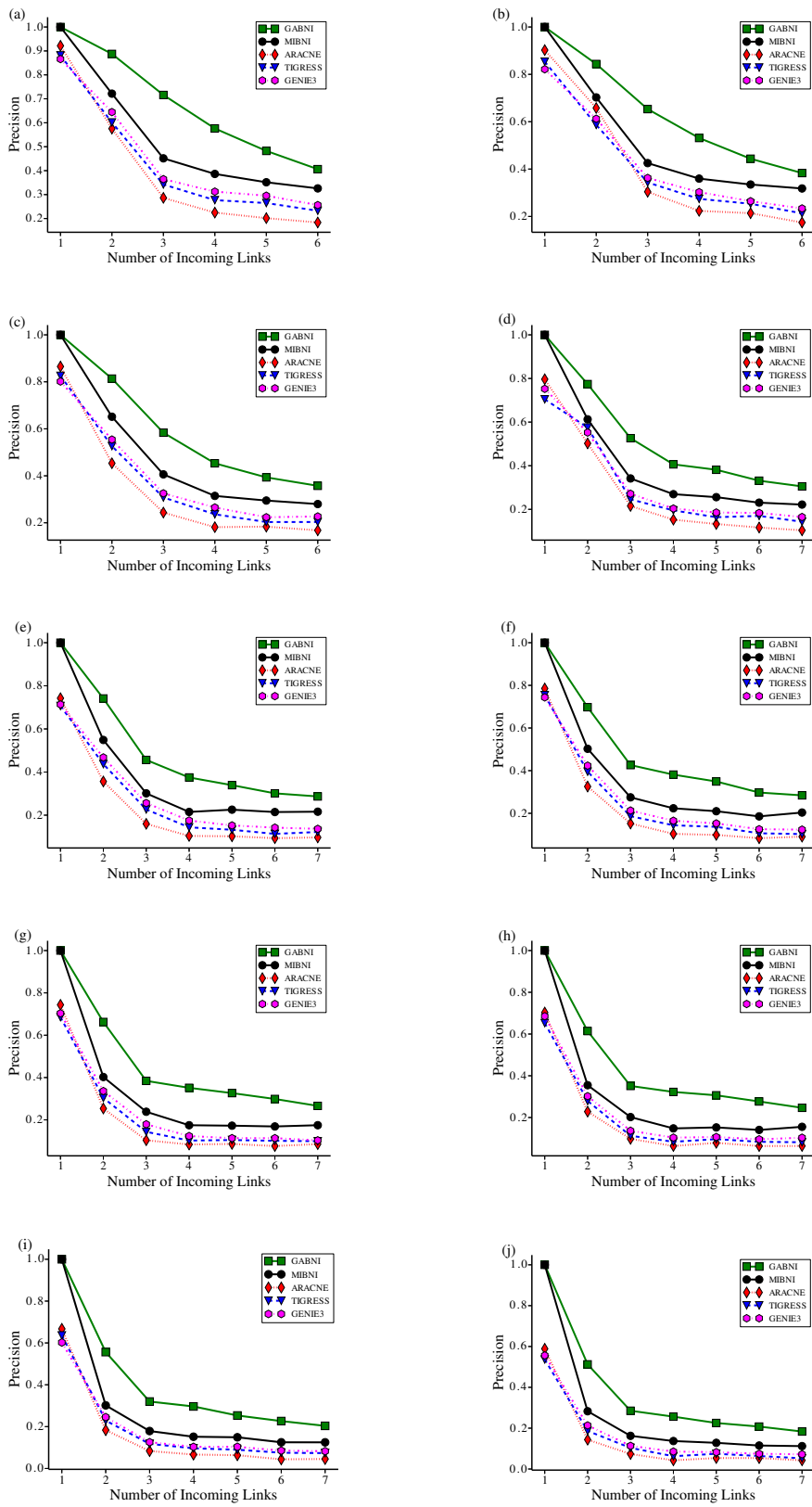
## Supplementary Information

```

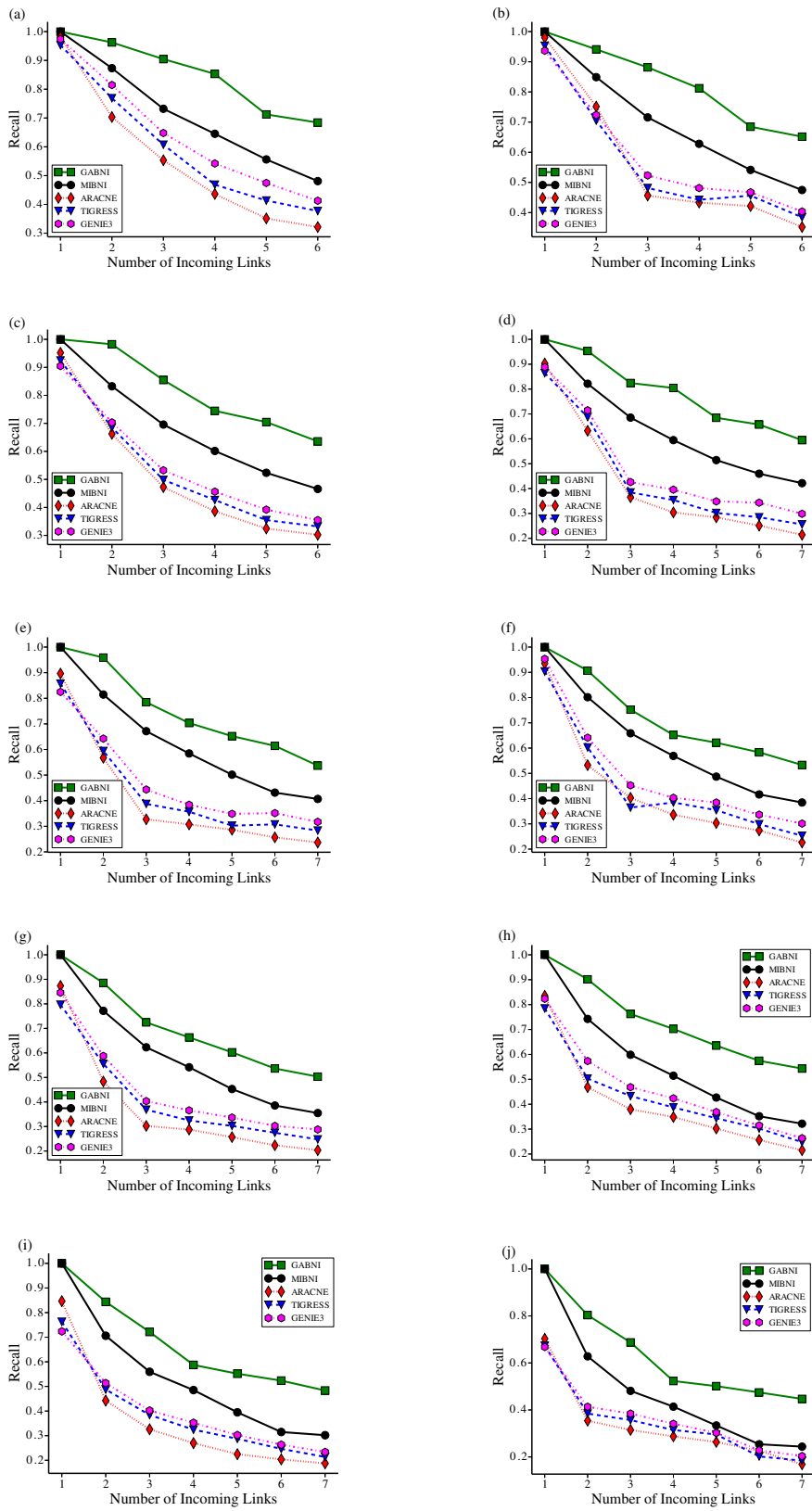
function [V, A] = createRBN_BarabasiAlbert(N, e, d)
    //N : The desired number of nodes
    //e : The number of initial nodes
    //d : The number of interactions to be added at each step
    //V, A : A set of nodes V and a set of links A of the resulting network
    generated by the Barabasi-Albert model
    //randNumber(0, 1) : Returns a real number chosen from 0 to 1 uniformly
    at random
    V ← {0, 1, ..., e - 1};
    A ← ∅
    for i := 0 to e - 2 do
        for j := i + 1 to e - 1 do
            if (randNumber(0, 1) < 0.5) then
                A ← A ∪ {(i, j)};
            else
                A ← A ∪ {(j, i)};
            end if
        end for
    end for
    for i := e to N - 1 do
        for j := 0 to d - 1 do
            repeat
                v = selection(V); //v is chosen with a probability proportional
                to its degree.
                if (randNumber(0, 1) < 0.5) then
                    vSrc ← i, vDst ← v;
                else
                    vSrc ← v, vDst ← i;
                end if
                until ((vSrc, vDst) ∉ A);
                A ← A ∪ {(vSrc, vDst)};
            end for
            V ← V ∪ {i};
        end for
    return [V, A]
end function

```

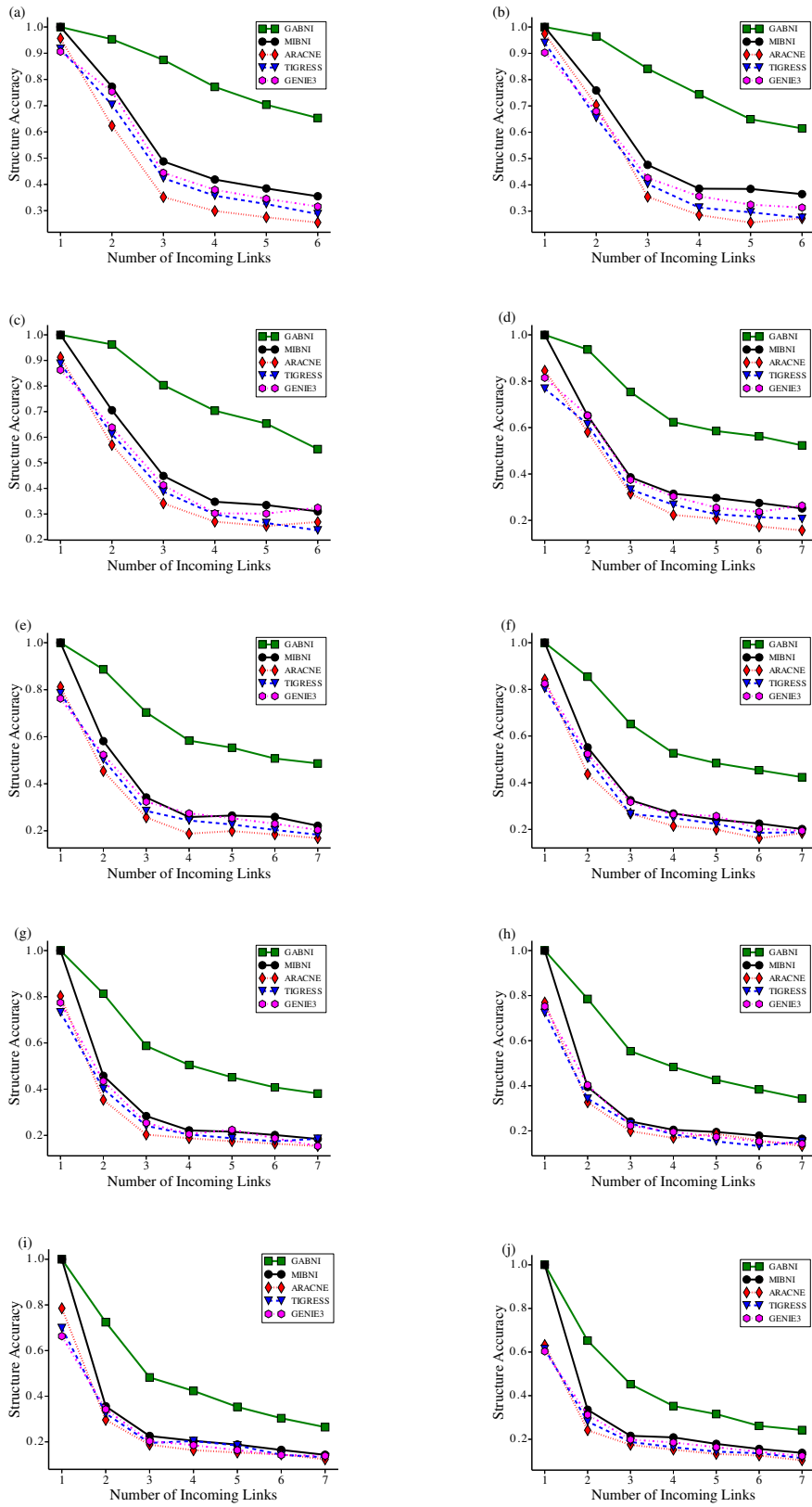
**Fig. S1: Pseudo-code of the Barabasi-Albert model used in our simulation.** The desired number of nodes ( $N$ ), the number of nodes in the seed network ( $e$ ), and the number of edges to be added at each iteration ( $d$ ) are given as parameters. A small seed network  $G(V, A)$  consisting of  $V = \{v_1, v_2, \dots, v_e\}$  and  $A = \{(v_i, v_j) | i, j = 1, 2, \dots, e, \text{ and } i \neq j\}$  is randomly specified. At each iteration, a node  $v$  and  $d$  different interactions are newly inserted into the graph where the probability of connecting a new node and an existing node is proportional to the connectivity of the latter node. This process is repeated until  $|V| = N$  and a resultant network is returned as the output.



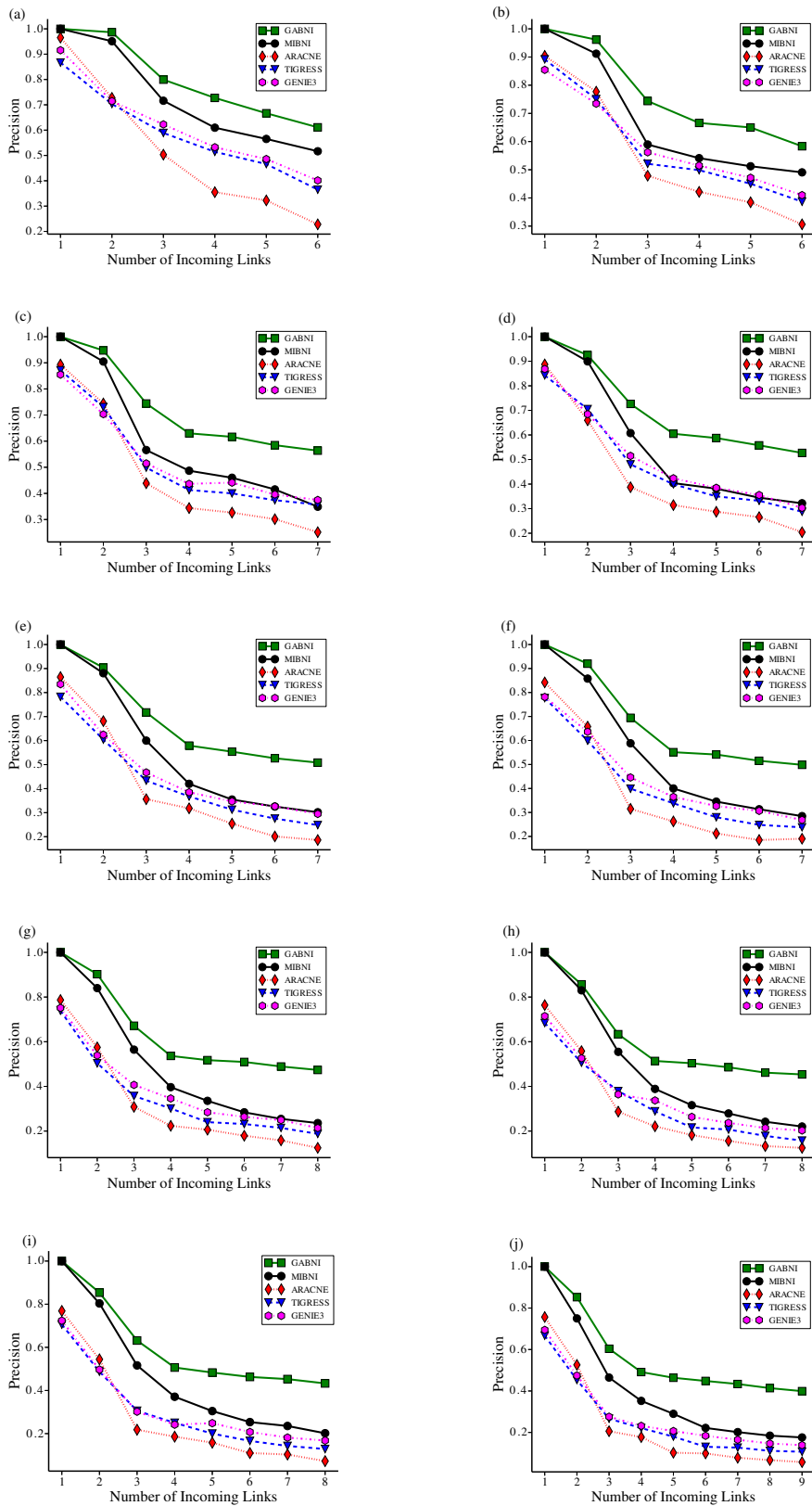
**Fig. S2: Comparison of precision between GABNI and other methods in GNW random networks.** (a)-(j) Precision results of networks with  $|V| = 10, 20, \dots, 100$ , respectively. In each subfigure, 30 GNW random networks were examined.



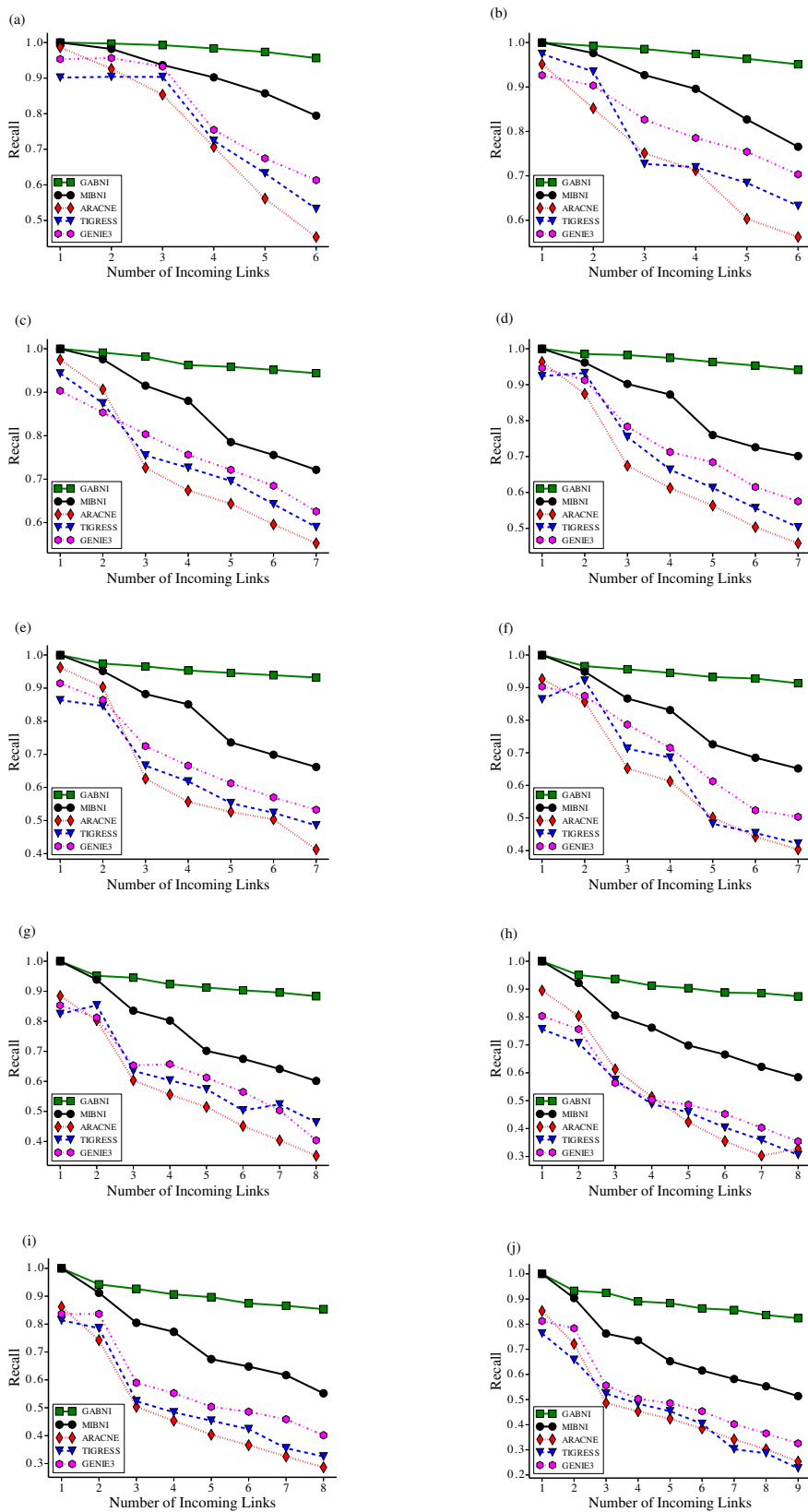
**Fig. S3: Comparison of recall between GABNI and other methods in GNW random networks.** (a)-(j) Recall results of networks with  $|V| = 10, 20, \dots, 100$ , respectively. In each subfigure, 30 GNW random networks were examined.



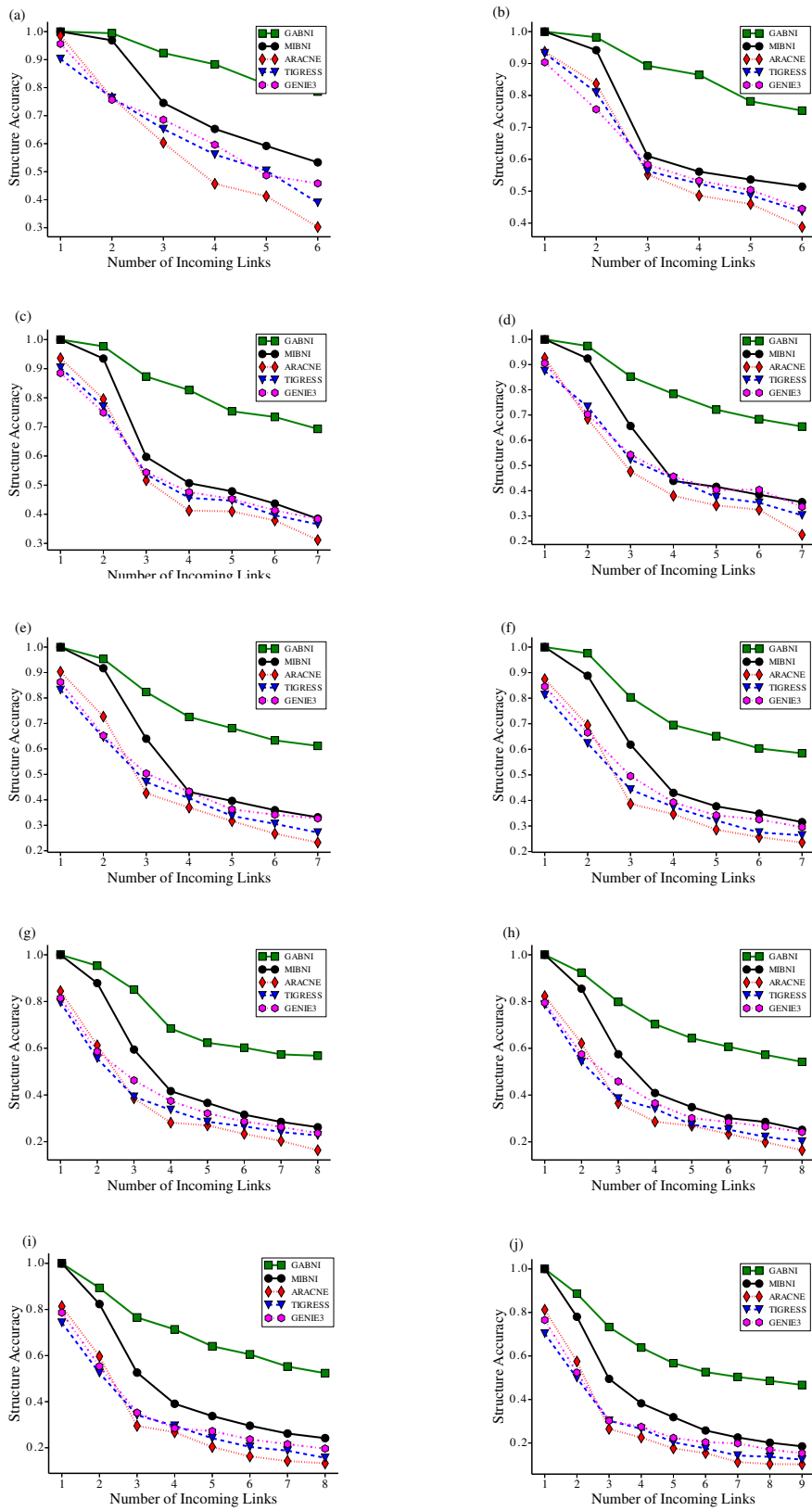
**Fig. S4: Comparison of structural accuracy between GABNI and other methods in GNW random networks. (a)-(j) Structural accuracy results of networks with  $|V| = 10, 20, \dots, 100$ , respectively. In each subfigure, 30 GNW random networks were examined.**



**Fig. S5: Comparison of precision between GABNI and other methods in BA random networks.** (a)-(j) Precision results of networks with  $|V| = 10, 20, \dots, 100$ , respectively. In each subfigure, 30 BA random networks were examined.

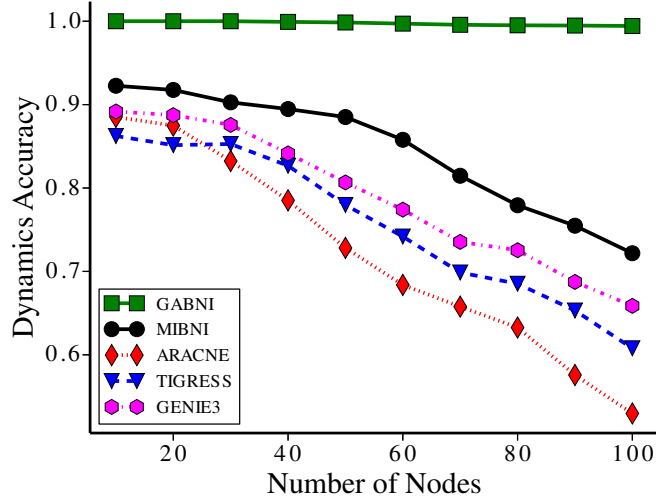


**Fig. S6: Comparison of recall between GABNI and other methods in BA random networks.** (a)-(j) Recall results of networks with  $|V| = 10, 20, \dots, 100$ , respectively. In each subfigure, 30 BA random networks were examined.

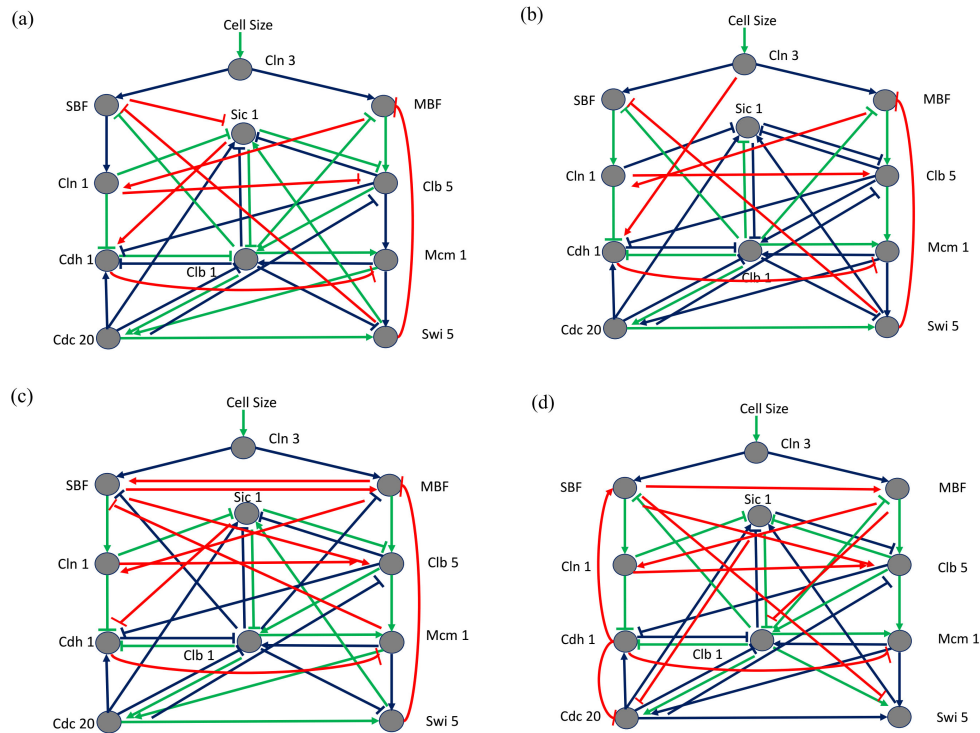


**Fig. S7: Comparison of structural accuracy between GABNI and other methods in BA random networks. (a)-(j) Structural accuracy results of networks with  $|V| = 10, 20, \dots, 100$ , respectively. In each subfigure, 30 BA random networks were examined.**





**Fig. S8: Comparison of dynamic accuracies between GABNI and other methods in GNW random networks.** A total of 300 GNW random networks with different network sizes ( $|V| = 10, 20, \dots, 100$ ) were used as target networks. For each point, 30 GNW random networks were examined. The dynamic accuracy of GABNI was 1.0 (perfect accuracy) for  $|V| \leq 30$  and 0.99 for the other sizes, whereas that of other methods sharply decreased as the network size increased.



**Fig. S9: Inference performance of MIBNI, ARACNE, GENIE3, and TIGRESS with respect to budding yeast cell cycle network.** The green, red, and blue interactions denote true positive, false positive, and false negative predictions, respectively. (a) Inference results of MIBNI. The results showed 14 true positives, 7 false positives, and 15 false negatives. The structural and dynamics accuracies were 0.8240 and 0.9600, respectively. (b) Inference results of ARACNE. The results showed 11 true positives, 6 false positives, and 18 false negatives. The structural and dynamics accuracies were 0.8125 and 0.9600, respectively. (c) Inference results of GENIE3. The results showed 14 true positives, 9 false positives, and 15 false negatives. The structural and dynamics accuracies were 0.8080 and 0.9800, respectively. (d) Inference results of TIGRESS. The results showed 14 true positives, 9 false positives, and 15 false negatives. The structural and dynamics accuracies were 0.8080 and 0.9500, respectively.

**Table. S1:** The number of nodes and the range of the number of edges in GNW random networks.

Number of nodes	Range of the number of edges
10	27-29
20	40-58
30	55-76
40	75-86
50	93-114
60	100-137
70	134-150
80	151-169
90	177-194
100	188-204

**Table. S2:** Boolean gene expression dataset of the budding yeast cell cycle.

Time	Cln3	MBF	SBF	Cln1	Cdh1	Swi5	Cdc20	Clb5	Sic1	Clb1	Mcm1	Phase
1	1	0	0	0	1	0	0	0	1	0	0	Start
2	0	1	1	0	1	0	0	0	1	0	0	G1
3	0	1	1	1	1	0	0	0	1	0	0	G1
4	0	1	1	1	0	0	0	0	0	0	0	G1
5	0	1	1	1	0	0	0	1	0	0	0	S
6	0	1	1	1	0	0	0	1	0	1	1	G2
7	0	0	0	1	0	0	1	1	0	1	1	M
8	0	0	0	0	0	1	1	0	0	1	1	M
9	0	0	0	0	0	1	1	0	1	1	1	M
10	0	0	0	0	0	1	1	0	1	0	1	M
11	0	0	0	0	1	1	1	0	1	0	0	M
12	0	0	0	0	1	1	0	0	1	0	0	M
13	0	0	0	0	1	0	0	0	1	0	0	G1

**Table. S3:** The structural and dynamical accuracies in inferring the Budding yeast cell cycle network from a noisy gene expression.

Inference methods	TP	FP	FN	Structural accuracy	Dynamical accuracy	Noise level
GABNI	18	3	11	0.8843	1.0000	0% Noise
MIBNI	14	7	15	0.8240	0.9600	
ARACNE	11	6	18	0.8125	0.9600	
GENIE3	14	9	15	0.8080	0.9800	
TIGRESS	14	9	15	0.8080	0.9500	
GABNI	15	11	14	0.8000	0.9800	5% Noise
MIBNI	11	20	18	0.7076	0.8800	
ARACNE	8	20	21	0.6870	0.8700	
GENIE3	10	28	19	0.6356	0.9100	
TIGRESS	9	29	20	0.6230	0.8500	
GABNI	14	13	15	0.7741	0.9700	10% Noise
MIBNI	9	23	20	0.6692	0.8600	
ARACNE	7	34	22	0.5757	0.8500	
GENIE3	9	31	20	0.6076	0.9000	
TIGRESS	8	32	21	0.5954	0.8600	